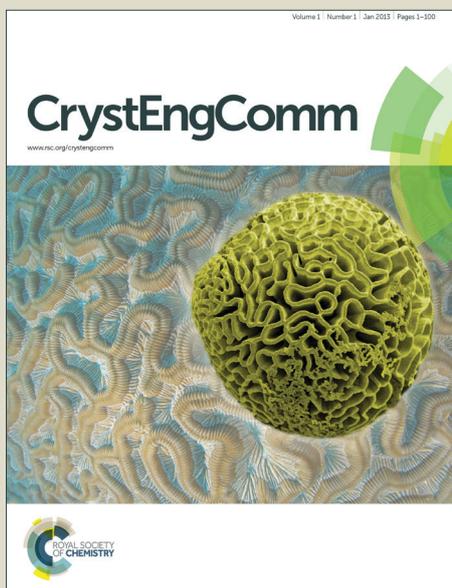


CrystEngComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Virtual hydrate screening and coformer selection for improved relative humidity stability

Yuriy A. Abramov

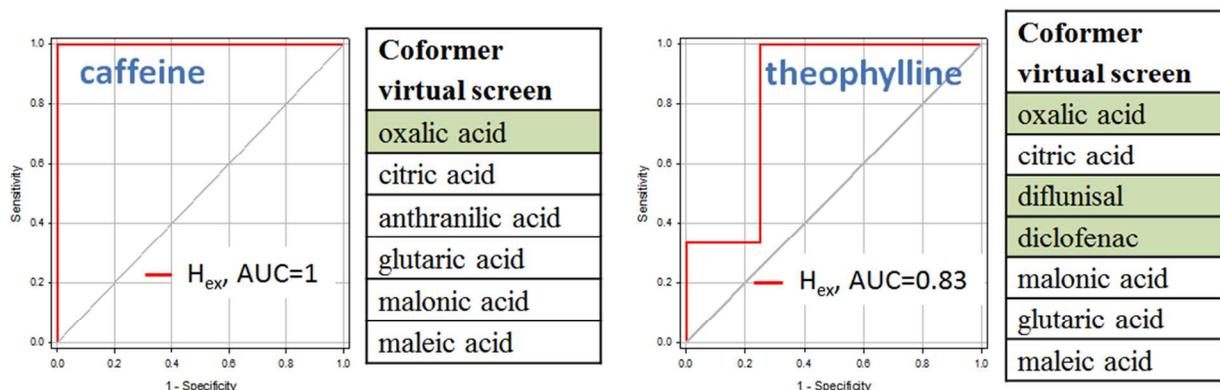
Pfizer Global Research and Development, Groton, Connecticut 06340, USA; E-mail: yuriy.a.abramov@pfizer.com

Abstract

Drug formulations of anhydrous solid forms are generally preferred over hydrated forms. This is due to the risks of low exposure and unacceptable physical and chemical stability in comparison with anhydrous formulations. The purpose of the current study was to determine which descriptors can be most efficiently applied to virtual screening in order to provide answers to the following questions: 1) what is the propensity to form a solid state hydrate of a pharmaceutical compound, and 2) in regards to cocrystalline formulation, which coformer would provide for the highest stability with respect to relative humidity (RH) conditions? A number of properties of different complexity were tested to provide answers to these questions, including COSMO-RS excess free energy G_{ex} and enthalpy H_{ex} of hydration of the compound in amorphous state; octanol-water partition coefficient clogP ; polar surface area TPSA ; different combinations of molecular H-bond donor and acceptor counts; an excess enthalpy of API (active pharmaceutical ingredient)-coformer mixing; and coformer solubilities. It was demonstrated that the G_{ex} property provides the most efficient way of virtual screening of hydration propensity of solid pharmaceutical compounds. It was also demonstrated that a virtual coformer screening based on the API coformer miscibility, as measured by the COSMO-RS H_{ex} property, may be efficiently used to guide the experimental selection of cofomers which have an increased probability of cocrystallization and provide the highest RH stability.

Graphical abstract

The descriptors were determined, which can be most efficiently applied to virtual screening in order to provide answers to the following questions: 1) what is the propensity to form a solid state hydrate of a pharmaceutical compound, and 2) which coformer would provide for the highest stability with respect to relative humidity conditions?



1. Introduction

Crystalline hydrates of pharmaceutical compounds appear to be a widespread phenomenon: while they represent only about 6.5% of all organic entries archived in the Cambridge Structure Database (CSD),¹⁻³ recent analysis of 245 polymorph screens performed at SSCI (<http://www.ssci-inc.com>) indicated that 38% of all found forms were hydrates.⁴ In the pharmaceutical industry, hydrates may have a high impact on the performance of an API mainly due a dramatic decrease of thermodynamic aqueous solubility,⁵ which can affect bioavailability of the drug compound.⁶ In addition, solid hydrate formation may lead to increased rate of chemical degradation^{7, 8} and may effect physical properties, such as particle size⁹ and physical stability due to hydrate-anhydrous form interconversion during process and/or storage conditions.^{10, 11} As a result, a drug formulation from a hydrate solid form is not generally

desirable; hydrates represent only a few percent of the total number of approved APIs.¹² Therefore, gauging the propensity of pharmaceutical compounds to form hydrates at design stage of drug discovery would allow mitigating the risk of unexpected solid hydrate formation during lead development and PK studies in particular.

The thermodynamic stability of solid hydrates is driven by the water activity in the crystallization or storage media exceeding a critical value.¹³ This relationship is explored in experimental hydrate screening approaches, which are carried out using such techniques as crystallization from aqueous solvent systems, aqueous slurries, or exposure to high relative humidity (RH) conditions.^{14, 15} The latter approach is typically used in physical and chemical stability testing of anhydrous APIs under various temperature (up to 60 °C) and humidity conditions (up to 85%) for up to 8 weeks.¹⁶ In case an undesired stable hydrate of API is found, one of the proposed strategies is to perform coformer screening to prepare a cocrystal that, unlike the API, is physically stable at high RH conditions.¹⁷⁻²¹ However, typically all the hydrate screening experiments in the pharmaceutical industry are carried out only at the drug development stage. Therefore, an ability to rationalize and predict hydration propensity of an API or its cocrystal is quite important and may bring a high value to the drug design at early stages of drug design.

There have been several attempts to rationalize solid hydrate formation.^{3, 22-28} Factors important for hydration of organic crystals have been extensively studied based on the CSD database mining. In an early study,²² Desiraju proposed that formation of solid hydrate is favored when the number of H-bond acceptors (HBA) exceeds the number of donor (HBD) groups in

molecules. This consideration suggests that water molecules are predominantly compensating donor-acceptor imbalance in the molecular crystal. However, the observation was not confirmed by a more elaborated study by Infantes et al³ conducted later. In that study, a simple count of HBD and HBA atom types is replaced by an average (potential) number of donor (NDavg) or acceptor (NAavg) H-bonds per atom, which was estimated by counting the total number of contacts, donor or acceptor, per atom in the training sample of 34,770 crystal structures and dividing this by the number of observations of this atom type. A positive correlation with a frequency of hydrate formation was found for such parameters as SDA and DAdiff. SDA is defined as the sum of NAavg and NDavg estimates over all atoms in the molecule. DAdiff is defined as an absolute difference of the sum of NDavg and the sum of NAavg bonds in the molecule. In addition it was found that increasing molecular polar surface is correlated with increased hydrate formation.³ The latter observation agrees with the result of an earlier CSD survey²⁴ which found that probability of organic molecules crystallizing as hydrates increases with increasing number of polar chemical groups, especially with ionic charge. The CSD was also systematically searched for hydrate formation of N-based cations and pharmaceutically acceptable anions.²⁵ The most pronounced observation found was a trend towards reduced hydration in pyridinium carboxylate salts. These models, built on the basis of the CSD database, are very useful tools in helping guide scientists to understand hydration in molecular crystals. However, the database does not provide information on organic compounds that completed hydrate screening and yet did not form solid hydrates. The fact that there is no hydrate form reported for a specific compound in the CSD does not mean that it cannot be made. As a result, the models built on the basis of CSD are inherently biased towards a prediction of hydrates rather than anhydrous crystals.

Rationalization of stoichiometric solid hydrates of organic molecules has been successfully undertaken based on prediction and analysis of lattice energy landscape.^{23, 26-28} However, though the crystal structure prediction field has considerably progressed over the last decade²⁹ and became a useful tool for complementing experimental pharmaceutical solid form selection,³⁰ it is still a computationally expensive approach and cannot be routinely applied towards virtual hydrate screening.

The purpose of the current study is to determine which computational descriptors can be most efficiently applied to virtual screening in order to provide answers to the following questions: 1) what is the propensity to form a solid state hydrate of a pharmaceutical compound, and 2) in regards to cocrystalline formulation, which coformer would provide for the highest stability with respect to relative humidity (RH) conditions?

In a recent study,³¹ it was demonstrated that a conductor-like screening model for real solvents (COSMO-RS)^{32, 33} is a powerful approach to perform virtual screening of solvents for solid desolvation and coformers for cocrystallization. In the current study, the COSMO-RS approach is specifically extended towards virtual hydrate screening of organic molecules. The performance of the COSMO-RS approach is compared to virtual screening results based on such properties as octanol-water partition coefficient clogP; molecular topological polar surface area TPSA; as well as the ratio, sum and absolute difference of number of H-bond donors and acceptors in the molecule. It is also demonstrated that the COSMO-RS approach can be

successfully applied to virtual coformer screening for an increased cocrystal resistance to hydration.

2. Approach and methods

2.1 Approach

In the recent study,³¹ we demonstrated that COSMO-RS fluid-phase thermodynamics computations can be used for accurate and efficient virtual screening of coformers for cocrystallization and solvents for solid desolvation. The hit enrichment was based on the excess enthalpy (H_{ex}) property that describes miscibility of conformer or solvent with API in an amorphous (supercooled liquid) state:

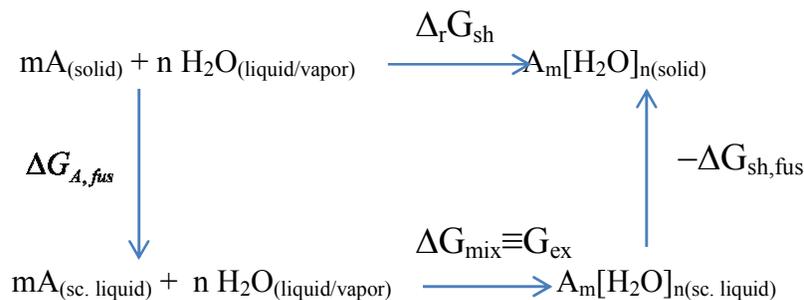
$$H_{\text{ex}} = H_{\text{AB}} - x_m H_{\text{pure,A}} - x_n H_{\text{pure,B}}, \quad (1)$$

Here H_{pure} and H_{AB} represent the molar enthalpies in the pure reference state and in the $m:n$ mixture, with mole fractions $x_m = m/(m + n)$ and $x_n = n/(m + n)$. Solvents or coformers, which have the highest/lowest probability to form solid solvates or cocrystals, are determined by the corresponding lowest/highest H_{ex} values.

The stabilizing or destabilizing contributions to H_{ex} of the lattice energies of each component is currently impossible to take into account in an efficient and reasonably accurate manner. Omitting lattice energy contribution remains the major source of error in the COSMO-RS predictions. Nevertheless, the enrichment of coformer or solvent selection based on amorphous (supercooled liquid) simulations was demonstrated to be quite good.³¹

The application of the excess enthalpy (H_{ex}) rather than excess free energy (G_{ex}) of mixing for virtual screening of coformers was dictated by the solid state nature of all components of cocrystallization reaction (an API, coformer and cocrystal), for which an entropic contribution should not be that pronounced. In the case of solid solvates, application of an enthalpic H_{ex} descriptor for solvent selection for desolvation seems to be less justified based on the fact that solvent is a liquid in the pure state. In that case, an entropic contribution should account for loss of solvent flexibility due to solidification during solid solvate formation. Nevertheless, applicability of H_{ex} parameter to solvent selection for desolvation was supported by good hit enrichment results found during multiple studies.³¹ However, water is a special case of solvent system, and is traditionally associated with a strong entropic contribution to various properties. For example, transfer of a hydrophobic molecule from octanol to water is entropically unfavorable, which is consistent with a “classical” definition of hydrophobic effect.³⁴ Therefore, in addition to the H_{ex} descriptor, the excess free energy property, G_{ex} , was tested to rank solid hydrate formation of pharmaceutical molecules.

Solid hydrate formation may be described by a thermodynamic cycle presented on Scheme 1.



$$\Delta_r G_{sh} = \Delta G_{ex} + \Delta \Delta G_{fus} = \Delta G_{ex} + (\Delta G_{A, fus} - \Delta G_{sh, fus})$$

Scheme 1. Thermodynamic cycle of solid hydrate formation. Here $\Delta_r G_{sh}$ is the free energy of solid hydrate formation from a crystalline compound A and water molecule(s); G_{ex} is an excess free energy of hydration of the compound A in amorphous (supercooled liquid) phase; $\Delta G_{A,fus}$ and $\Delta G_{sh,fus}$ are free energies of fusion of the crystals of pure compound A and its solid hydrate, respectively.

Two limiting cases can be distinguished. In the first scenario, the hydration propensity is dominated by water miscibility in a supercooled phase of the compound: $|G_{ex}| \gg |\Delta\Delta G_{fus}|$. This condition justifies virtual hydrate screening based on the G_{ex} property alone. In the second limiting case, lattice energy contributions to solid hydrate formation are dominant: $|\Delta\Delta G_{fus}| \gg |G_{ex}|$. In that case, virtual screening based on G_{ex} property will fail.

While experimental $\Delta G_{A,fus}$ values may be readily available, the $\Delta G_{sh,fus}$ property is unknown and practically impossible to measure due to dehydration effects prior to melting. Nevertheless, one may speculate that it will be difficult to counterbalance a very strong lattice energy of API (very high $\Delta G_{A,fus}$ value) with even stronger packing of the corresponding solid hydrate. This gives rise to an assumption that a limiting scenario of the high $\Delta\Delta G_{fus}$ contribution to $\Delta_r G_{sh}$ may be approximated by the observation of an extremely high value of $\Delta G_{A,fus}$. Therefore, whenever experimental fusion properties (T_m and ΔH_{fus}) were available, $\Delta G_{A,fus}$ was considered as an additional parameter describing solid hydrate formation propensity.

In addition to the computationally advanced H_{ex} and G_{ex} properties, the following simplified descriptors were also tested for the virtual hydrate propensity screening: octanol-water partition coefficient clogP , molecular topological polar surface area TPSA, as well as ratio (RDA), sum (SDA) and absolute difference (DDA) of the molecular HBD and HBA counts. clogP is a very popular descriptor in drug discovery field that describes molecular hydrophobicity in a simplistic 2-dimensional way and mimics molecular hydration free energy contribution to aqueous solubility.^{35, 36} TPSA, RDA, SDA and DDA are simplified versions of the corresponding properties that were previously found to correlate with hydrate formation frequency based on the CSD surveys.^{3, 22}

2.2 Computational methods

Excess enthalpies, H_{ex} , and free energies, G_{ex} , were calculated by the COSMOtherm software (Eckert F, Klamt A, 2014, COSMOtherm, version C3.0, release 14.01). For the compounds that were not available in the COSMObase database, multiple conformations were generated using mixed MCMM/Low-Mode algorithm and OPLS_2005 forcefield as implemented in MacroModel program (MacroModel; Schrodinger, LLC, New York, 2011; <http://www.schrodinger.com>).

In order to generate screening charge densities for COSMOtherm calculations, the generated conformations were further optimized in aqueous media by the Turbomole package (TURBOMOLE, TURBOMOLE V6.3 2011, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007; available

from <http://www.turbomole.com>) using the BP86 density functional³⁷⁻³⁹ with a TZVP⁴⁰ basis set (BP-TZVP-COSMO level of theory).

H-bonding donor and acceptor counts and TPSA descriptor were calculated from the molecular structures by MOE software package (MOE (The Molecular Operating Environment) Version 2013.08, Chemical Computing Group Inc. Available from: <http://www.chemcomp.com>).

Free energy of fusions at ambient temperature (T) were generated from experimental melting points, T_m , and heats of fusion, ΔH_{fus} , according to the equation: $\Delta G_{fus} = \Delta H_{fus} (T/T_m) \ln(T_m/T)$. It was demonstrated previously⁴¹ that for drug-like molecules this approach is superior to the most popular one of $\Delta G_{fus} = \Delta H_{fus} (1 - T/T_m)$.

2.3 Data set

In this study, 61 experimental observations for single component crystalline compounds (in a neutral or zwitterionic form) were used for *in silico* hydrate formation tests (Table 1). The majority of these compounds are pharmaceutical APIs. From the 61 experimental examples, 41 observations of hydration were taken from different literature sources of hydrated APIs. The most challenging task was to find the remaining 20 cases of compounds not forming solid hydrates. From the set of compounds that did not form hydrates, the most reliable results represent 14 compounds which completed an experimental hydrate formation screening (Pfizer compounds A-J, torcetrapib, crizotinib, indomethacin¹⁴ and chlorthiazide¹⁴). The RH stability profiles of caffeine and theophylline cocrystals were taken from literature sources.^{17, 18, 42, 43}

3. Results and discussions

3.1 Hydrate virtual screening of pharmaceutical compounds

Results of virtual hydrate screening of the APIs (compounds) based on H_{ex} , G_{ex} , clogP, TPSA, RDA, SDA and DDA properties are compared with the experimental observations in Tables 1 and 2. Overall performance of the *in silico* models for hydrate formation screening was estimated by receiver operator characteristic (ROC) curves (Fig. 1). An ROC curve plots the sensitivity (a true positive rate equal to the number of true positive predictions/total number of positive observations) versus 1-specificity (a false positive rate equal to the number of false-positive predictions/total number of negative observations) for a binary classifier system (hydrate formation result) as its discrimination threshold (descriptor cutoff) is continuously varied. The area under the curve (AUC) measures the overall performance of the model. Predictions with higher AUCs are generally better and should always be higher than 0.5, indicating that the model is better than random selection. A step-like ROC curve with AUC of 1 represents a perfect prediction.

Table 1. Pharmaceutical compounds and related properties for hydrate virtual screening test. Excess energies G_{ex} and H_{ex} and free energy of fusion ΔG_{fus} are presented in kcal/mol units. Zwitterionic compounds are indicated by “zw” index. TPSA is a topological surface area; SDA, DDA and RDA are sum, an absolute difference and ratio of the donor and acceptor counts in the molecule, respectively.

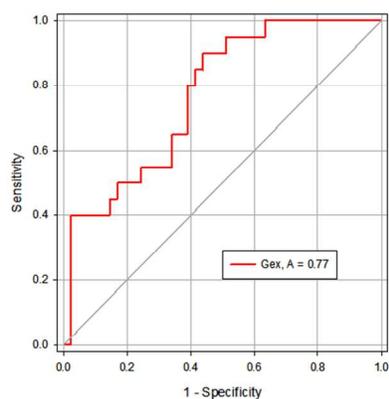
Compound	Hydrate	G_{ex}	H_{ex}	clogP	TPSA	SDA	DDA	RDA	ΔG_{fus}	References
ROY	N	0.579	0.494	1.48	108.76	5	1	1.50		44
thiophanate-ethyl	N	0.568	0.603	2.46	164.9	8	0	1.00		45
torcetrapib	N	0.540	0.290	7.55	59.08	4	4	0.00	1.38	
thiophanate-methyl	N	0.524	0.597	1.40	164.9	8	0	1.00		46
Compound A	N	0.519	0.386	4.31	74.86	4	2	0.33		
Compound B	N	0.494	0.272	4.70	53.33	3	3	0.00		
Compound C	N	0.413	0.350	2.18	94.47	5	3	0.25	3.34	
Compound D	N	0.397	0.183	2.45	152.61	10	6	0.25		

sulfathiazole	N	0.326	0.350	0.73	121.7	6	0	1.00	1.90	47
Compound E	N	0.282	-0.150	5.00	63.33	4	2	0.33		
indomethacin	N	0.258	-0.114	4.18	68.53	5	3	0.25	2.26	14
Compound F	N	0.231	-0.157	1.80	91.02	5	3	0.25	3.45	
Compound G	N	0.223	0.226	0.49	122.2	6	0	1.00	2.89	
chlorothiazide	N	0.195	0.301	-0.29	135.45	7	1	0.75		14
diclofenac	N	0.193	-0.424	4.73	49.33	4	0	1.00	2.69	42
2,5-hydrobenzoic acid	N	0.189	-0.200	1.37	37.3	3	1	0.50		28
Compound H	N	0.176	-0.082	2.54	99.36	6	2	0.50	3.23	
Compound I	N	0.169	-0.171	2.21	103.54	8	2	0.60	2.29	
Compound J	N	0.152	-0.271	3.40	59	5	3	0.25		
crizotinib	N	0.093	-0.885	4.29	77.99	7	1	0.75	8.28	
niclosamide	Y	0.606	0.220	4.35	92.47	5	1	0.67	3.03	14
glutethimide	Y	0.362	0.183	1.99	46.17	3	1	0.50		6
nifedipine	Y	0.362	0.213	3.13	107.77	5	3	0.25	2.31	48
neotame	Y	0.357	0.143	0.67	104.73	8	2	0.60		49
aripiprazole	Y	0.341	-0.173	5.31	44.81	4	2	0.33		50
dacomitinib	Y	0.330	-0.268	5.90	79.38	7	3	0.40		
piroxicam (zw)	Y	0.289	0.162	-1.46	112.06	3	1	2.00	2.60	51
estradiol	Y	0.275	0.052	3.78	40.46	4	0	1.00		52
axitinib	Y	0.271	0.013	3.33	95.97	6	2	0.50	3.70	53, 54
tranilast	Y	0.264	-0.075	3.99	84.86	7	3	0.40		15
nitrofurantoin	Y	0.257	0.403	-3.50	119.61	6	0	1.00		55
etoricoxib	Y	0.257	-0.258	2.35	68.3	4	4	0.00		56
calcipotriol	Y	0.256	-0.055	5.27	60.69	6	0	1.00		57
dasatinib	Y	0.244	-0.035	2.53	134.75	9	3	0.50		58
2,4-hydrobenzoic acid	Y	0.200	-0.222	1.85	37.3	3	1	0.50		28
caffeine	Y	0.199	-0.128	-0.04	61.82	3	3	0.00	1.60	14, 15
carbamazepine	Y	0.189	-0.268	2.38	46.33	3	1	2.00	1.90	14, 15
Compound I	Y	0.170	-0.294	0.74	87.46	5	1	0.67		59
fluprednisolone	Y	0.166	-0.166	1.99	94.83	8	2	0.60		60
fluconazole	Y	0.159	-0.406	-0.44	81.65	4	2	0.33		15
cefaclor (zw)	Y	0.158	-0.083	-1.64	138.03	10	2	0.67		61
acetaminophen	Y	0.148	-0.004	0.49		4	0	1.00	1.78	62
diflunisal	Y	0.126	-0.554	4.40	57.53	5	1	0.67	2.57	42, 63
theophylline	Y	0.123	-0.097	-0.03	72.68	4	2	0.33	2.34	14, 15
cimetidine	Y	0.109	-0.182	0.38	114.19	7	1	0.75	2.49	14
flunisolide	Y	0.105	-0.147	2.41	93.06	8	4	0.33		64
levofloxacin	Y	0.082	-0.329	-0.51	75.01	6	4	0.20		65
isonicotinamide	Y	0.070	-0.202	-0.21	55.98	4	0	1.00		66
5-azauracil	Y	0.035	0.040	-1.87	65.6	5	1	1.50		67

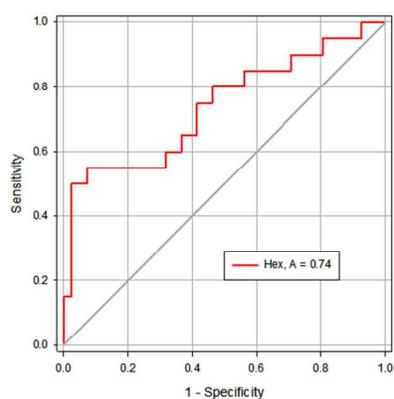
dextrose	Y	0.011	-0.159	-2.21	110.38	11	1	0.83		68
lactose	Y	0.001	-0.203	-4.40	189.53	19	3	0.73		15
lamivudine	Y	-0.016	-0.326	-1.46	115.67	8	2	0.60		69
acyclovir	Y	-0.030	-0.187	-2.42	119.05	9	1	0.80		70
cefadroxil (zw)	Y	-0.124	-0.179	-0.20	134.44	12	0	1.00		71
citric acid	Y	-0.168	-4.173	-2.00	132.13	11	3	0.57		15
ampicillin (zw)	Y	-0.174	-0.525	-1.72	142.48	9	1	0.80		68
L-phenylalanine	Y	-0.190	-0.684	-2.21	67.77	5	1	1.50		72
gallic acid	Y	-0.210	-0.792	0.43	97.99	9	1	0.80		73
cephradine (zw)	Y	-0.266	-0.782	0.70	114.21	10	0	1.00		74
norfloxacin (zw)	Y	-0.575	-1.656	1.81	63.24	7	1	0.75	2.93	75
ciprofloxacin	Y	-0.578	-1.667	1.86	63.24	7	1	0.75		76

Table 2. Summary of performance results of hydration virtual screening based on the different descriptors. Positive direction specifies whether low or high values of the descriptor are expected to provide an increased probability of anhydrous form.

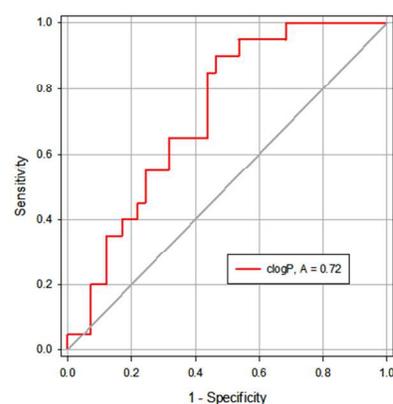
Descriptor	Positive direction	AUC
G_{ex}	High	0.77
H_{ex}	High	0.74
clogP	High	0.72
RDA	High	<0.5
DDA	Low	<0.50
SDA	Low	0.57
TPSA	Low	0.57



a



b



c

Fig. 1 ROC curves of virtual hydration test based on G_{ex} (a), H_{ex} (b) and clogP (c) properties.

It was found (Table 2) that the simplified descriptors, such as polar surface area descriptor TPSA and different combinations of HBA and HBD counts, provide a poor prediction of hydration propensities of pharmaceutical organic molecules. However, the overall performance of models based on G_{ex} , H_{ex} and clogP properties is quite good as demonstrated by the AUC values greater than 0.7 (Table 2, Fig.1). A detailed side-by-side comparison of the ROC curves demonstrates a clear advantage of excess energy descriptors over clogP property (Fig. 1). Screening based on the clogP descriptor provides the slowest enrichment rate (true positive rate growth) among these methods (Fig. 1). Though the fastest enrichment is achieved by virtual screening based on excess enthalpy property, the G_{ex} descriptor demonstrates a noticeably more balanced enrichment throughout the screening resulting in the highest AUC value of 0.77. Optimal cutoff values calculated from G_{ex} and H_{ex} ROC plots are 0.38 kcal/mol and 0.22 kcal/mol, respectively.

The poorer behavior of the clogP descriptor for hydrate formation screen can be accounted for by the following drawbacks. First of all, in contrast to G_{ex} and H_{ex} , the octanol-water partition of organic drug-like molecules does not adequately represent any of the steps of the thermodynamic cycle presented in Scheme 1. In addition, clogP is a 2-dimensional property that does not take into account conformational flexibility of the molecules as some other more advanced *in silico* partition methods do.⁷⁷ The simplistic nature of TPSA, RDA, SDA and DDA descriptors may also account for the failure for these parameters to accurately describe the propensity of hydration formation.

The ΔG_{fus} properties are available for 21 compounds from the whole data set. The values of free energy of fusion ranges from 1.4 kcal/mol to 8.3 kcal/mol (Fig. 2). An extremely high ΔG_{fus} value for crizotinib puts it into the second limiting group discussed above with $|\Delta\Delta G_{\text{fus}}| \gg |G_{\text{ex}}|$. This accounts for the poor performance of G_{ex} property for the prediction of hydration propensity of crizotinib (Table 1).

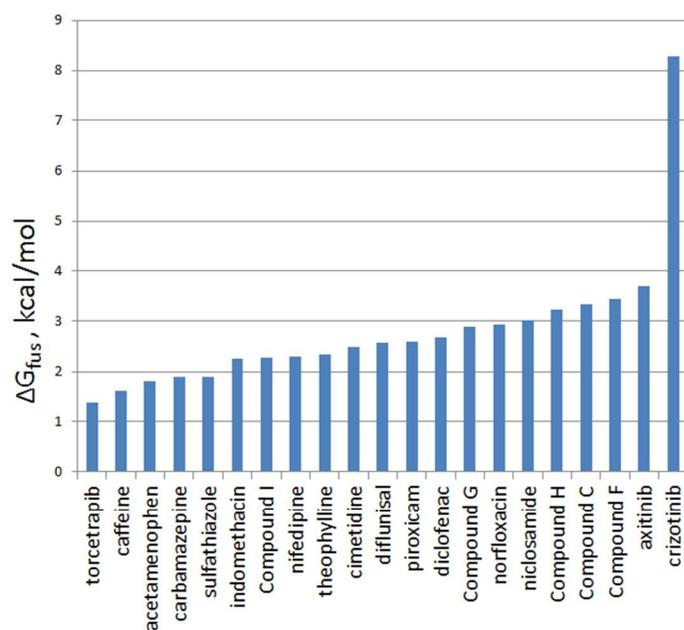


Fig. 2 ΔG_{fus} property distribution.

3.2 *In silico* coformer screening for an improved stability at high RH

It has been shown previously that cocrystallization of APIs that form both hydrate and anhydrous solid forms (such as carbamazepine, caffeine and theophylline) may lead to stabilization against hydrate formation.¹⁷⁻¹⁹ It was also recently demonstrated^{21, 78} that in spite of improved resistance of some cocrystals to hydration, a partial dissociation occurs under all humidity conditions even for the cocrystals that do not form hydrates. This is believed to be driven by widely differing aqueous solubilities and hygroscopicities of two cocrystal formers, as is typically designed to

improve aqueous dissolution rate of the API (compound). Mechanistically, at increased relative humidity, the highly soluble coformer interacts with water molecules and gets “washed off” the cocrystal forming a separate amorphous or liquid state.⁷⁸ For the hydrate-forming APIs, such dissociation eventually leads to crystalline hydrate formation even at a microscopic level.⁷⁸ This mechanism of hydrate formation suggests that cocrystallization of hydrate forming APIs may typically provide a kinetic rather than thermodynamic hydration stability at a high relative humidity. Nevertheless, an increased resistance to hydrate formation is still a task of great importance in pharmaceutical industry.

Based on the described mechanism of cocrystal dissociation, it is reasonable to conclude that cocrystal resistance to hydration should increase with a stronger coformer-API interaction and lower coformer solubility and hygroscopicity. While experimental coformer solubility values are typically available, their hygroscopicities depend on particle properties and are not easy to measure or predict. Therefore, in this study we tested virtual screening models that are based on API-coformer miscibility (interaction) in a supercooled liquid state as measured by excess enthalpy H_{ex} ,³¹ as well as on experimental coformer aqueous solubility values, S_{cof} .⁷⁹⁻⁸¹ A miscibility of water with the amorphous stoichiometric cocrystal was estimated based on predicted G_{ex} value, which was also tested as a descriptor for the virtual coformer screening.

For testing of the approaches, the experimental hydration screening results for cocrystals of caffeine and theophylline APIs were taken from literature sources.^{17, 18, 42, 43} In those studies, all caffeine^{18, 43} and theophylline^{17, 42, 43} cocrystals were initially grown in water-free media and afterwards were subjected to different relative humidity (RH) conditions for time periods of up to

7 weeks. Such procedures are quite relevant to hydrate stability testing of solid forms in the pharmaceutical industry,¹⁶ as they have practical implications for processing, formulation, packaging and storage conditions.

Virtual screenings results are compared with the experimental observations after up to 7 weeks at 98% RH in Tables 3 and 4. The quality of the predictions based on cocrystal formers miscibility H_{ex} , cofomers solubilities S_{cof} and water cocrystal miscibility descriptor G_{ex} , were measured by the ROC curves presented on Fig. 3 and 4.

Table 3. Properties of caffeine cocrystals. Here S_{cof} is an experimental aqueous solubility of cofomer; H_{ex} is an excess enthalpy of cofomer and caffeine mixing in a soopercooled liquid state; G_{ex} is an excess free energy of mixing of cocrystal with 1 water molecule per caffeine in a soopercooled liquid state. Cocrystals that were found to be stable to hydrate formation after 7 weeks at 98% RH are highlighted in grey.

Cocrystal	S_{cof} , M	H_{ex} , kcal/mol	G_{ex} , kcal/mol	References
caffeine/oxalic acid 2:1	1.3	-3.63	0.079	18, 79, 81
caffeine/citric acid 1:1	1.99	-2.89	0.027	43, 80
caffeine/anthranilic acid	0.025	-2.07	0.175	80, 82
caffeine/glutaric acid 1:1	10.68	-1.474	0.073	18, 79, 81
caffeine/malonic acid 2:1	15.3	-1.351	0.108	18, 79, 81
caffeine/maleic acid 1:1	6.86	-1.173	0.074	18, 79, 81

Table 4. Properties of theophylline cocrystals. Here S_{cof} is experimental aqueous solubility of coformer; H_{ex} is an excess enthalpy of coformer and caffeine mixing in a supercooled liquid state; G_{ex} is an excess free energy of mixing of cocrystal with 1 water molecule per theophylline in a supercooled liquid state. Cocrystals that were found to be stable to hydrate formation after 7 weeks at 98% RH are highlighted in grey. Cocrystals of theophylline with diclofenac and diflunisal were not tested at 98% RH but were found to be physically stable at 75% and 100% RH for up to 2 months.⁴²

Cocrystal	S_{cof} , M	H_{ex} , kcal/mol	G_{ex} , kcal/mol	References
theophylline/oxalic acid 2:1	1.3	-3.738	0.04	17,76, 78
theophylline/citric acid 1:1	1.99	-2.213	0.05	43, 80
theophylline/diflunisal 1:1	0.0155	-1.808	0.164	42
theophylline/diclofenac 1:1	0.0052	-1.614	0.182	42
theophylline/malonic acid 1:1	15.3	-1.116	0.04	17,76, 78
theophylline/glutaric acid 1:1	10.68	-1.018	0.048	17,76, 78
theophylline/maleic acid 1:1	6.86	-0.84	0.049	17,76, 78

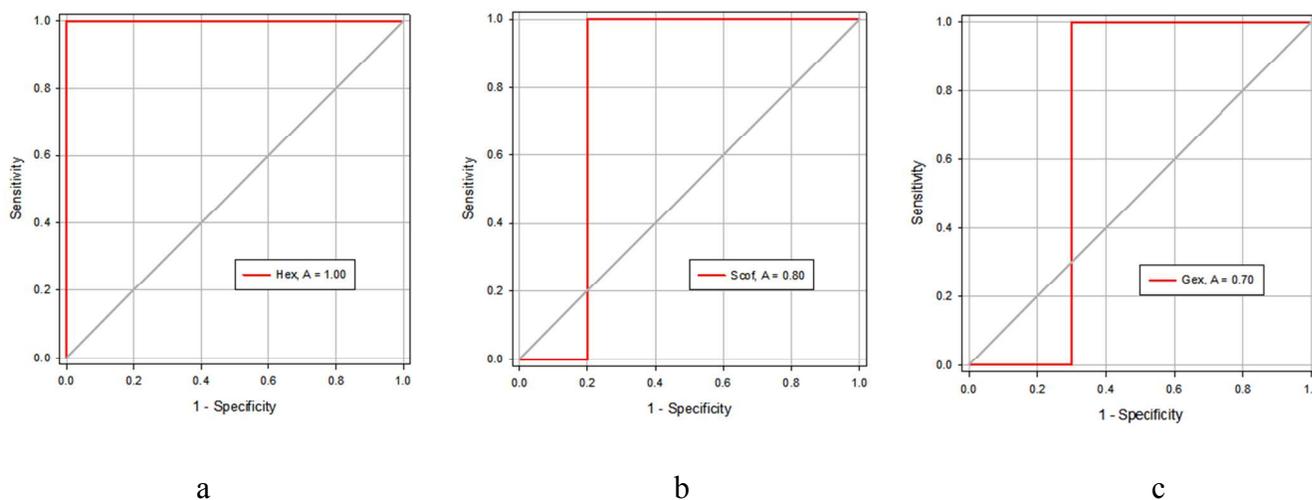


Fig. 3 ROC curves of virtual coformer screening on hydration resistance of caffeine cocrystals based on (a) cocrystal formers miscibility H_{ex} , (b) cofomers solubilities S_{cof} and (c) water cocrystal miscibility descriptor G_{ex} .

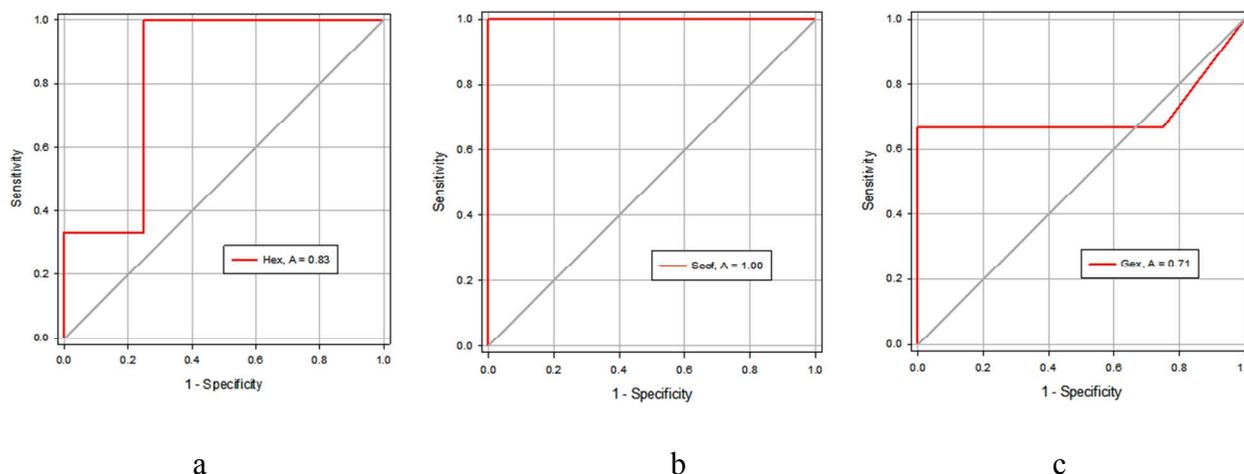


Fig. 4 ROC curves of virtual coformer screening on hydration resistance of theophylline cocrystals based on (a) cocrystal formers miscibility H_{ex} , (b) coformers solubilities, S_{cof} and (c) water cocrystal miscibility descriptor G_{ex} .

The best and most consistent virtual screening results were based on the H_{ex} property measuring API-coformer miscibility. An agreement with experimental outcomes is excellent (AUC=1.0) for the hydrate virtual screening of caffeine cocrystals. There is only one misranking (citric acid) in the case of theophylline virtual coformer screening, resulting in an excellent hit enrichment (AUC of 0.83). Virtual screenings based on coformer solubility values appear to be comparably successful, resulting in AUC values for caffeine and theophylline cocrystals of 0.80 and 1.0, respectively. The poorest enrichment was found for the virtual screening based on excess free energy G_{ex} of water mixing with a stoichiometric cocrystal in a supercooled liquid phase. The corresponding AUC values are 0.70 in case of caffeine and 0.71 in case of theophylline.

For the systems considered in this study, there is a stronger interaction (a higher miscibility) between cocrystal formers as measured by a lower H_{ex} property that plays the dominant role in

the improved resistance to hydration. As was shown previously,³¹ the virtual screening hits, determined by the lowest H_{ex} values, also represent the optimal coformer selection for cocrystallization. The S_{cof} parameter is also found to be a very important one, which agrees with the previous opinion^{20, 21, 78} that hydration stability of cocrystals is governed by a lower coformer solubility property. However, in this study the coformer solubility is an experimental rather than computational parameter which may not be available for a broader set of cofomers (for example, from GRAS (<http://www.fda.gov/Food/IngredientsPackagingLabeling/GRAS/>) or EAFUS (<http://www.accessdata.fda.gov/scripts/fcn/fcnNavigation.cfm?rpt=eafusListing>) molecular libraries).

4. Conclusions

The goal of the current study was to determine which descriptors can be most efficiently applied for virtual screening to provide answers to the following questions: 1) what is the hydration propensity of a solid pharmaceutical compound, and 2) which conformer(s) would provide the highest resistance to a pharmaceutical cocrystal hydration at high RH conditions? A number of computational properties of different complexity were tested to provide the answers to the first question, including excess free energy G_{ex} and enthalpy H_{ex} of hydration of the compound in a supercooled liquid state; octanol-water partition coefficient clogP ; topological polar surface area; and different combinations of the donor and acceptor counts in the molecule (such as ratio, sum and an absolute difference). It is demonstrated that COSMO-RS theory, as implemented in the *COSMOtherm* software, offers the most efficient way to test the hydration propensity of a solid pharmaceutical compound. This is achieved by calculation of the G_{ex} property that describes the miscibility of a water molecule in an amorphous pharmaceutical compound prior to crystallization. As a result, the G_{ex} property may be used as an additional parameter in a multi-

parameter optimization of pharmaceutical compounds in order to mitigate the risk of unexpected solid hydrate formation during PK studies.

It was also demonstrated that for the caffeine and theophylline cocrystals considered in this study, there is a stronger interaction between cocrystal formers, as measured by a lower H_{ex} property, that plays the dominant role in the improved resistance to API hydration at 98% RH conditions. As a result, a virtual cofomer screening based on the H_{ex} property may be used to guide the experimental selection of cofomers that have an increased probability of cocrystallization and provide an improved RH stability.

Acknowledgments

The author would like to thank his colleagues Mr. Brian Samas, Mr. Andrew J. Jensen, Dr. Klimentina Pencheva and Ms. Shawn LaCasse for providing information on hydrate screening of Pfizer compounds for this study. The author is also grateful to Mr. Brian Samas for valuable discussions.

References

1. F. H. Allen, *Acta Crystallographica Section B: Structural Science*, 2002, **58**, 380-388.
2. J. van de Streek and S. Motherwell, *CrystEngComm*, 2007, **9**, 55-64.
3. L. Infantes, L. Fábián and W. S. Motherwell, *CrystEngComm*, 2007, **9**, 65-71.
4. G. P. Stahly, *Crystal growth & design*, 2007, **7**, 1007-1026.
5. M. Pudipeddi and A. Serajuddin, *Journal of pharmaceutical sciences*, 2005, **94**, 929-939.
6. R. K. Khankari and D. J. Grant, *Thermochimica Acta*, 1995, **248**, 61-79.
7. H. Mimura, K. Gato, S. Kitamura, T. Kitagawa and S. Kohda, *Chemical and pharmaceutical bulletin*, 2002, **50**, 766-770.
8. S. Ohtake and E. Shalaev, *Journal of pharmaceutical sciences*, 2013, **102**, 1139-1154.
9. S. N. Black, A. Phillips and C. I. Scott, *Organic Process Research & Development*, 2008, **13**, 78-83.
10. M. P. Feth, N. Nagel, B. Baumgartner, M. Bröckelmann, D. Rigal, B. Otto, M. Spitzenberg, M. Schulz, B. Becker and F. Fischer, *European Journal of Pharmaceutical Sciences*, 2011, **42**, 116-129.
11. A. D. Gift, P. E. Luner, L. Luedeman and L. S. Taylor, *Journal of pharmaceutical sciences*, 2009, **98**, 4670-4683.

12. B. Kratochvíl, in *Glassy, Amorphous and Nano-Crystalline Materials*, Springer, 2011, pp. 129-140.
13. D. J.W. Grant and T. Higuchi, *Solubility Behavior of Organic Compounds*, Wiley, New York, 1990.
14. A. Sistla, Y. Wu, P. Khamphavong and J. Liu, *Pharmaceutical development and technology*, 2011, **16**, 102-109.
15. Y. Cui and E. Yao, *Journal of pharmaceutical sciences*, 2008, **97**, 2730-2744.
16. X.-Q. Chen, M. D. Antman, C. Gesenberg and O. S. Gudmundsson, *The AAPS journal*, 2006, **8**, E402-E408.
17. A. V. Trask, W. Motherwell and W. Jones, *International journal of pharmaceutics*, 2006, **320**, 114-123.
18. A. V. Trask, W. S. Motherwell and W. Jones, *Crystal growth & design*, 2005, **5**, 1013-1021.
19. S. L. Childs, N. Rodríguez-Hornedo, L. S. Reddy, A. Jayasankar, C. Maheshwari, L. McCausland, R. Shipplett and B. C. Stahly, *CrystEngComm*, 2008, **10**, 856-864.
20. A. Jayasankar, L. Roy and N. Rodríguez-Hornedo, *Journal of pharmaceutical sciences*, 2010, **99**, 3977-3985.
21. M. D. Eddleston, N. Madusanka and W. Jones, *Journal of pharmaceutical sciences*, 2014, **103**, 2865-2870.
22. G. R. Desiraju, *J. Chem. Soc., Chem. Commun.*, 1991, 426-428.
23. B. P. van Eijck and J. Kroon, *Acta Crystallographica Section B: Structural Science*, 2000, **56**, 535-542.
24. L. Infantes, J. Chisholm and S. Motherwell, *CrystEngComm*, 2003, **5**, 480-486.
25. D. A. Haynes, W. Jones and W. S. Motherwell, *Cryst. Eng. Comm*, 2005, **7**, 342-345.
26. A. T. Hulme and S. L. Price, *Journal of Chemical Theory and Computation*, 2007, **3**, 1597-1608.
27. K. M. Anderson, G. M. Day, M. J. Paterson, P. Byrne, N. Clarke and J. W. Steed, *Angewandte Chemie*, 2008, **120**, 1074-1078.
28. D. E. Braun, P. G. Karamertzanis and S. L. Price, *Chemical Communications*, 2011, **47**, 5443-5445.
29. G. M. Day, *Crystallography Reviews*, 2011, **17**, 3-52.
30. Y. A. Abramov, *Organic Process Research & Development*, 2012, **17**, 472-485.
31. Y. A. Abramov, C. Loschen and A. Klamt, *Journal of pharmaceutical sciences*, 2012, **101**, 3687-3697.
32. A. Klamt, *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, Elsevier, 2005.
33. F. Eckert and A. Klamt, *AIChE Journal*, 2002, **48**, 369-385.
34. P. W. Snyder, J. Mecinović, D. T. Moustakas, S. W. Thomas, M. Harder, E. T. Mack, M. R. Lockett, A. Héroux, W. Sherman and G. M. Whitesides, *Proceedings of the National Academy of Sciences*, 2011, **108**, 17889-17894.
35. N. Jain and S. H. Yalkowsky, *Journal of pharmaceutical sciences*, 2001, **90**, 234-252.
36. C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Advanced drug delivery reviews*, 2012, **64**, 4-17.
37. A. D. Becke, *Physical Review A*, 1988, **38**, 3098-3100.
38. J. P. Perdew, *Physical Review B*, 1986, **33**, 8822-8824.
39. J. P. Perdew, *Physical Review B*, 1986, **34**, 7406(E).
40. A. Schäfer, C. Huber and R. Ahlrichs, *The Journal of Chemical Physics*, 1994, **100**, 5829-5835.
41. Y. A. Abramov and K. Pencheva, in *Chemical Engineering in the Pharmaceutical Industry: R&D to Manufacturing*, ed. D.J am Ende, Wiley, 2010, 477-490.
42. A. O. Surov, A. P. Voronin, A. N. Manin, N. G. Manin, L. G. Kuzmina, A. V. Churakov and G. L. Perlovich, *Molecular pharmaceutics*, 2014, **11**, 3707-3715.
43. S. Karki, T. Friščić, W. Jones and W. S. Motherwell, *Molecular pharmaceutics*, 2007, **4**, 347-354.
44. L. Yu, *Accounts of chemical research*, 2010, **43**, 1257-1266.

45. E. Nauha, A. Ojala, M. Nissinen and H. Saxell, *CrystEngComm*, 2011, **13**, 4956-4964.
46. E. Nauha, H. Saxell, M. Nissinen, E. Kolehmainen, A. Schäfer and R. Schlecker, *CrystEngComm*, 2009, **11**, 2536-2547.
47. A. L. Bingham, D. S. Hughes, M. B. Hursthouse, R. W. Lancaster, S. Tavener and T. L. Threlfall, *Chemical Communications*, 2001, 603-604.
48. M. R. Caira, Y. Robbertse, J. J. Bergh, M. Song and M. M. De Villiers, *Journal of pharmaceutical sciences*, 2003, **92**, 2519-2533.
49. Z. Dong, J. S. Salsbury, D. Zhou, E. J. Munson, S. A. Schroeder, I. Prakash, S. Vyazovkin, C. A. Wight and D. J. Grant, *Journal of pharmaceutical sciences*, 2002, **91**, 1423-1431.
50. D. E. Braun, T. Gelbrich, V. Kahlenberg, R. Tessadri, J. Wieser and U. J. Griesser, *Crystal Growth and Design*, 2008, **9**, 1054-1065.
51. A. R. Sheth, D. Zhou, F. X. Muller and D. J. Grant, *Journal of pharmaceutical sciences*, 2004, **93**, 3013-3026.
52. B. Busetta, *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry*, 1972, **28**, 560-567.
53. B. P. Chekal, A. M. Campeta, Y. A. Abramov, N. Feeder, P. P. Glynn, R. W. McLaughlin, P. A. Meenan and R. A. Singer, *Organic Process Research & Development*, 2009, **13**, 1327-1337.
54. A. M. Campeta, B. P. Chekal, Y. A. Abramov, P. A. Meenan, M. J. Henson, B. Shi, R. A. Singer and K. R. Horspool, *Journal of pharmaceutical sciences*, 2010, **99**, 3874-3886.
55. V. R. Vangala, P. S. Chow and R. B. Tan, *Crystal Growth & Design*, 2012, **12**, 5925-5938.
56. C. R. Dalton, S. D. Clas, J. Singh, K. Khougaz and R. Bilbeisi, *Journal of pharmaceutical sciences*, 2006, **95**, 56-69.
57. S. Larsen, E. T. Hansen, L. Hoffmeyer and N. Rastrup-Andersen, *Acta Crystallographica Section C: Crystal Structure Communications*, 1993, **49**, 618-621.
58. S. Roy, R. Quiñones and A. J. Matzger, *Crystal growth & design*, 2012, **12**, 2122-2126.
59. N. Variankaval, C. Lee, J. Xu, R. Calabria, N. Tsou and R. Ball, *Organic process research & development*, 2007, **11**, 229-236.
60. J. K. Haleblan, R. T. Koda and J. A. Biles, *Journal of pharmaceutical sciences*, 1971, **60**, 1485-1488.
61. H. Martinez, S. R. Byrn and R. R. Pfeiffer, *Pharmaceutical research*, 1990, **7**, 147-153.
62. P. A. McGregor, D. R. Allan, S. Parsons and C. R. Pulham, *Journal of pharmaceutical sciences*, 2002, **91**, 1308-1311.
63. L. K. Hansen, G. L. Perlovich and A. Bauer-Brandl, *Acta Crystallographica Section E: Structure Reports Online*, 2001, **57**, o477-o479.
64. M. Bartolomei, *Journal of pharmaceutical and biomedical analysis*, 2000, **24**, 81-93.
65. E. M. Gorman, B. Samas and E. J. Munson, *Journal of pharmaceutical sciences*, 2012, **101**, 3319-3330.
66. N. B. Báthori, A. Lemmerer, G. A. Venter, S. A. Bourne and M. R. Caira, *Crystal Growth & Design*, 2010, **11**, 75-87.
67. B. Potter, R. Palmer and B. Chowdhry, *New Journal of chemistry*, 1999, **23**, 117-122.
68. H. B. Liu, Y. Chen and X. C. Zhang, *Journal of pharmaceutical sciences*, 2007, **96**, 927-934.
69. A. Bhattacharya, B. N. Roy, G. P. Singh, D. Srivastava and A. K. Mukherjee, *Acta Crystallographica Section C: Crystal Structure Communications*, 2010, **66**, o329-o333.
70. G. I. Birnbaum, M. Cygler and D. Shugar, *Canadian journal of chemistry*, 1984, **62**, 2646-2652.
71. W. Shin and S. W. Cho, *Acta Crystallographica Section C: Crystal Structure Communications*, 1992, **48**, 1454-1456.
72. J. Lu, J. Wang, Z. Li and S. Rohani, *African Journal of Pharmacy and Pharmacology*, 2012, **6**, 269-277.

73. D. E. Braun, R. M. Bhardwaj, A. J. Florence, D. A. Tocher and S. L. Price, *Crystal growth & design*, 2012, **13**, 19-23.
74. S. R. Vippagunta, H. G. Brittain and D. J. Grant, *Advanced drug delivery reviews*, 2001, **48**, 3-26.
75. S. Roy, N. R. Goud, N. J. Babu, J. Iqbal, A. K. Kruthiventi and A. Nangia, *Crystal Growth and Design*, 2008, **8**, 4343-4346.
76. L. Mafra, S. r. M. Santos, R. e. Siegel, I. s. Alves, F. A. Almeida Paz, D. Dudenko and H. W. Spiess, *Journal of the American Chemical Society*, 2011, **134**, 71-74.
77. M. Shalaeva, G. Caron, Y. A. Abramov, T. N. O'Connell, M. S. Plummer, G. Yalamanchi, K. A. Farley, G. H. Goetz, L. Philippe and M. J. Shapiro, *Journal of medicinal chemistry*, 2013, **56**, 4870-4879.
78. M. D. Eddleston, R. Thakuria, B. J. Aldous and W. Jones, *Journal of pharmaceutical sciences*, 2014, **103**, 2859–2864.
79. S. L. Clegg and J. H. Seinfeld, *The Journal of Physical Chemistry A*, 2006, **110**, 5692-5717.
80. S. Yalkowsky and Y. He, *Handbook of Aqueous Solubility Data: An Extensive Compilation of Aqueous Solubility Data for Organic Compounds Extracted from the AQUASOL dATABASE*, 2003.
81. D. J. Good and N. r. Rodríguez-Hornedo, *Crystal Growth and Design*, 2009, **9**, 2252-2264.
82. N. Madusanka, M. D. Eddleston, M. Arhangelskis and W. Jones, *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, 2014, **70**, 72-80.